

CLAIMS:

1. A method (1) for classifying at least one audio signal (A) into at least one audio class (AC), the method (1) comprising the steps of:

- analyzing (10) said audio signal to extract at least one predetermined audio feature;
- 5 - performing (12) a frequency analysis on a set of values of said audio feature at different time instances;
- deriving (12) at least one further audio feature representing a temporal behavior of said audio feature based on said frequency analysis; and
- classifying (14) said audio signal based on said further audio feature.

10

2. The method as claimed in claim 1, wherein said at least one predetermined audio feature comprises at least one of the following audio features:

- root-mean-square (RMS) level;
- spectral centroid (S_f);
- 15 - bandwidth (B_f);
- zero-crossing rate (R_z);
- spectral roll-off frequency (f_r);
- band energy ratio (B_r);
- delta spectrum magnitude (f_d);
- 20 - pitch (T); and
- pitch strength (S).

3. The method as claimed in claim 1, wherein said predetermined audio feature comprises at least one mel-frequency cepstral coefficient (MFCC).

25

4. The method as claimed in claim 1, wherein said predetermined audio feature comprises at least one of the psycho-acoustic (PA) audio features loudness and sharpness.

5. The method as claimed in claim 1, wherein said deriving step comprises the steps of:

- calculating an average (DC) value of said set of values of said audio feature at different time instances;
- 5 - defining at least one frequency band;
- calculating the amount of energy within said frequency band from said frequency analysis; and
- defining said further audio feature as said amount of energy in dependence on said average (DC) value.

10

6. The method as claimed in claim 5, wherein at least one of the following modulation frequency bands are used in said parameterizing step:

- 1-2 Hz;
- 3-15 Hz; and
- 15 - 20-150 Hz;

7. The method as claimed in claim 1, wherein said at least one further audio feature is defined as at least one coefficient (C(m)) obtained by performing a discrete cosine transformation (DCT) on the result of said frequency analysis.

20

8. A system (20) for classifying at least one audio signal into at least one audio class, the system comprising:

- means (10) for analyzing said audio signal to extract at least one predetermined audio feature;
- 25 - means (12) for performing a frequency analysis on a set of values of said audio feature at different time instances;
- means (12) for deriving at least one further audio feature representing a temporal behavior of said audio feature based on said frequency analysis; and
- means (14) for classifying said audio signal based on said further audio feature.

30 9. A music system (2) comprising:

- means (24) for playing audio data from a medium (22); and
- a system (20) as claimed in claim 8 for classifying said audio data.

10. A multi-media system (3) comprising:
 - means (34) for playing audio data from a medium (32);
 - a system (20) as claimed in claim 8 for classifying said audio data;
- 5 - means (36) for displaying video data from a further medium (32);
 - means (38) for analyzing said video data; and
 - means (34) for combining the results obtained from analyzing (38) said video data with the results obtained from classifying (20) said audio data.
- 10 11. A signal (22, 32) comprising at least one further audio feature obtained by:
 - analyzing (10) an audio signal to extract at least one predetermined audio feature;
 - performing (12) a frequency analysis on a set of values of said audio feature at different time instances; and
 - deriving (12) said at least one further audio feature representing a temporal behavior of said audio feature based on said frequency analysis.
- 15 12. An audio receiver (2) comprising means to receive an audio signal and means to receive a signal (22, 32) comprising at least one further audio feature obtained by:
 - analyzing (10) said audio signal to extract at least one predetermined audio feature;
 - performing (12) a frequency analysis on a set of values of said audio feature at different time instances; and
 - deriving (12) said at least one further audio feature representing a temporal behavior of said audio feature based on said frequency analysis,
 - 20 the audio receiver (24) further comprising means (14) for classifying said audio signal based on said further audio feature.
 - 25